

IR2

by Sneha Report

Submission date: 04-May-2018 09:08AM (UTC+0530)

Submission ID: 958664939

File name: sneha_IR2_Report.pdf (163.12K)

Word count: 1636

Character count: 9004



¹⁰ Department of Information Technology
National Institute of Technology Karnataka, Surathkal.
May 2018

Project Report

On

¹ **Automatic Medical Image Annotation
and Content-based Image Retrieval
using Relevance Feedback**

Submitted by

Manali Ashish Shah (15IT218)
Ria Kulshrestha (15IT231)
Shivani Shrivastava (15IT243)
Sneha Suresh Patil (15IT245)

VI Sem B.Tech

Under the Guidance of

¹² **Dr. Sowmya Kamath S**
Dept. of Information Technology,
NITK, Surathkal

Abstract

This project implements classification of ¹medical images, content-based medical image retrieval, and relevance feedback method for image retrieval for enhancing image classification and hence retrieval performance. For image classification our system uses ⁷Convolutional Neural Network(CNN). For content-based image retrieval, our retrieval system uses the results from the soft¹max classifier used in our CNN architecture. To enhance the performance further we combine our image retrieval system with relevance feedback mechanism in CNN.

Contents

Abstract	i
1 Introduction	1
2 Literature Review	2
2.1 Background	2
2.2 Identified Gaps	2
2.3 Problem Statement	2
2.4 Objectives	3
3 Methodology and Framework	4
3.1 Network Architecture	4
3.2 Training method	4
4 Implementation	6
4.1 Work Done	6
4.2 Results and Analysis	6
4.3 Innovative Work	6
4.4 Details of each individual's work w.r.t. project tasks.	6
5 Conclusion and Future Work	7
References	8

1 Introduction

Image annotation is the task of attaching meta-data in form of caption or keywords to any image. The need for automatic annotation of medical images is ever-growing. Medical images like CT scans, MRIs, X-ray scans play a crucial role in modern medicine. They aid in early detection, diagnosis and treatment.

Content based Image Retrieval is the content(features) based search for similar images in a large database. Keyword based image retrieval when done, requires manual annotation which has many practical limitations due to the ever-growing size of the database. Also, the accuracy of keyword based retrieval system will be dependent on the precision of the annotation done manually, hence prone to more error. The content based image retrieval techniques addresses both these problems.

Content based image retrieval can be achieved using a CNN based model. CNN or ConvNet is a sub-class of Deep neural networks. They are complex feed-forward neural networks. They do not require feature extraction as they learn the features on their own while training. Also they have high accuracy and hence used very commonly for these problems(classification and recognition). A CNN model in conjunction with massive amount of data automates the annotation process as well as boosts the classification performance.

ImageClef 2009 dataset acts a benchmark for medical annotation. The image sets(train and test) are based on the IRMA project. The images of the dataset are fully classified radio graphs. It consists of 20,000 training images and 10,000 testing images categorized into 193 classes. The class label of the images are complete IRMA codes.

³ 2 Literature Review

2.1 Background

Outcomes of Literature Review

[8] employs the method of k-nearest neighbors to retrieve information. The retrieval considers the euclidean distance for fetching the information. It considers all the documents are relevant.

[5] proposes the method of using the decision trees for retrieval that considers relevance feedback. The edges linking the nodes of the decision trees are given weights that are trained as per the feedback from the user. The features that are found in relevant images are given more weights than the ones not found.

[6] uses the technique of center symmetric linear binary patterns to extract texture feature from the images. The method suits well for X-Ray images since it has contrasting bright foreground over dark background region.

2.2 Identified Gaps

The technique for k-nearest neighbors doesn't consider the relevance or non-relevance of the document. It assumes that all the documents are relevant.

The pruning of the decision tree will affect the ratio of relevance and non-relevance in the dataset and hence will affect performance of the model. The merging of the nodes will also affect the ratio and hence the results.

An image region is said to be flat if it has a nearly uniform intensity that is the variance of the intensity values within the region is very low. The LBP feature is not robust on flat image areas since it is based on intensity differences. It also ignores the actual intensity level at the location it is computed on.

2.3 Problem Statement

¹
Content-based medical image retrieval and relevance feedback method for image retrieval for enhancing image retrieval performance.

2.4 Objectives

- Performing image classification based on content of the images.
- Building an image retrieval system.
- Provide mechanism for relevance feedback.
- Improving classification performance based on the relevance feedback.

3 Methodology and Framework

3.1 Network Architecture

Our multi-layered Convolutional Neural Network¹ (CNN)¹¹ has 5 convolutional layers which are followed by 4 fully connected layers⁸. Each convolutional layer is followed by a ReLu activation layer which is in turn followed by a max pooling layer. The dimensions of each layer is as follows:

- Input convolutional layer: 224x224 input tensor with 32 filter of 5x5x3 and outputs a 112x112x32 tensor after ReLu and max pooling.
- Second convolutional layer: 112x112 input tensor with 64 filter of 5x5x32 and outputs a 56x56x64 tensor after ReLu and max pooling.
- Third convolutional layer: 56x56 input tensor with 128 filter of 5x5x64 and outputs a 28x28x128 tensor after ReLu and max pooling.
- Fourth convolutional layer: 28x28 input tensor with 256 filter of 5x5x128 and outputs a 14x14x256 tensor after ReLu and max pooling.
- Fifth convolutional layer: 14x14 input tensor with 256 filter of 5x5x256 and outputs a 7x7x256 tensor after ReLu and max pooling.

Fully Connected layers:

- First fully connected layer: 2048 nodes
- Second fully connected layer: 256 nodes
- Third fully connected layer: 64 nodes
- Fourth fully connected layer: 193 nodes

⁶ The output of the fourth fully connected layer is fed into the Softmax classifier.

3.2 Training method

Factors to be considered while training of the model include:

- Choice of optimizer: Adam Optimizer
- Choice of loss function: Cross-Entropy Loss
- Batch size: 100

- Number of epochs: 5
- Initialization of weights/ choice of checkpoint: Random initialization
- Learning Rate: 1e-4

⁵ The back propagation algorithm is used to compute the gradient with respect to the parameters of the model in order to achieve gradient based optimization, we have used an Adam optimizer.

⁴ Adam, an algorithm for first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order moments.

The Softmax classifier is used to finally predict the class label of the input images. It gives the probability of each class as the output. The target class has the highest probability. All the probabilities lie between 0-1 and the summation of probabilities of all the classes is equal to 1.

Cross Entropy is used to calculate the loss while training. ² Cross-entropy loss, or log loss, measures the performance of a classification model whose output is a probability value between 0 and 1. Cross-entropy loss increases as the predicted probability diverges from the actual label.

4 Implementation

4.1 Work Done

The project is developed using TensorFlow, which is a deep learning software library developed by Google. It is used for high performance numerical computation. The main advantage of using TensorFlow is its flexible architecture which allows easy deployment of computation across various platforms.

Firstly we re-organized the data so that each folder contains images of the class identified by its folder name. Then we create a file named label.txt which contains the list of all such folders.

After that we read the list of folders from label.txt and extract features like height, width, color space, channels, format, filename, class label along with the image itself and create a file of the format '.tfrecord', which is a simple record-oriented binary format commonly used in TensorFlow applications. It also splits the dataset into two sets i.e. training and testing. The data is shuffled so that there is a random distribution amongst the two sets.

The CNN is trained for 5 epochs. It is then used to classify the testing images. The results of the classification are then stored in a CSV file. The CSV file contains the predicted label and the image name.

For the retrieval system we use the CSV file as reference. When a user enters a particular class number then a list of all the images with the same class in the CSV file is retrieved.

For feedback we ask the user to enter the image name and the relevant class number. After receiving the feedback, we retrain the model by freezing the convolutional layers and only updating the weights of the fully connected layers.

4.2 Results and Analysis

4.3 Innovative Work

To be added later

4.4 Details of each individual's work w.r.t. project tasks.

To be added later

5 Conclusion and Future Work

To be added later

References

- [1] <http://www.imageclef.org/2009/medanno>
- [2] <https://www.sciencedirect.com/science/article/pii/S0925231208001355>
- [3] Ojala T, Pietikainen M, Maenpaa T: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24:971987, 2002
- [4] Latex Template for NITK, Adithya Bhat, Dept. of Information Technology, NITK, <https://github.com/adithyabhatkajake/latex-template-nitk>
- [5] MacArthur SD, Brodley CE, Shyu CR: Relevance feedback decision trees in content-based image retrieval. Proceedings of the IEEE Workshop on Content-Based Access of Image and Video Libraries: 6873, 2000.
- [6] Ko BC, Kim SH, Nam JY: X-ray image classification using random forests with local wavelet-based CS-local binary patterns. JDigit Imaging, 2011
- [7] Maria Tzelepi, Anastasios Tefas: Relevance Feedback in Deep Convolutional Neural Networks for Content Based Image Retrieval, 2016
- [8] Marius Muja, David G. Lowe:Fast Approximate Nearest Neighbors,2009 WITH AUTOMATIC ALGORITHM CONFIGURATION

ORIGINALITY REPORT

17%

SIMILARITY INDEX

9%

INTERNET SOURCES

12%

PUBLICATIONS

%

STUDENT PAPERS

PRIMARY SOURCES

1

www.theradiologydaily.com

Internet Source

4%

2

ml-cheatsheet.readthedocs.io

Internet Source

2%

3

edepot.wur.nl

Internet Source

2%

4

Wenliang Shao, Chunhong Zhang, Tingting Sun, Hang Li, Yang Ji, Xiaofeng Qiu.

"Hierarchical Bidirectional Long Short-Term Memory Networks for Chinese Messaging Spam Filtering", 2017 3rd International Conference on Big Data Computing and Communications (BIGCOM), 2017

Publication

1%

5

Ali El Housseini, Abdelmalek Toumi, Ali Khenchaf. "Deep Learning for target recognition from SAR images", 2017 Seminar on Detection Systems Architectures and Technologies (DAT), 2017

Publication

1%

6	"Face Verification Across Aging Based on Deep Convolutional Networks and Local Binary Patterns", Lecture Notes in Computer Science, 2015. Publication	1%
7	Maria Tzelepi, Anastasios Tefas. "Deep convolutional image retrieval: A general framework", Signal Processing: Image Communication, 2018 Publication	1%
8	"Advances in Multimedia Information Processing - PCM 2016", Springer Nature, 2016 Publication	1%
9	en.youscribe.com Internet Source	1%
10	zenodo.org Internet Source	1%
11	"Computer Vision – ECCV 2016", Springer Nature, 2016 Publication	1%
12	Kumar, Ajay, Ankit Goyal, Ashwani Kumar, Navneet Kumar Chaudhary, and S Sowmya Kamath. "Comparative evaluation of algorithms for effective data leakage detection", 2013 IEEE CONFERENCE ON INFORMATION AND COMMUNICATION TECHNOLOGIES, 2013. Publication	1%

13

Wang, Anran, Jiwen Lu, Jianfei Cai, Tat-Jen Cham, and Gang Wang. "Large-margin Multimodal Deep Learning for RGB-D Object Recognition", IEEE Transactions on Multimedia, 2015.

Publication

1%

14

Byoung Chul Ko. "Automatic medical image annotation and keyword-based image retrieval using relevance feedback", Journal of Digital Imaging, 12/23/2011

Publication

1%

Exclude quotes On

Exclude bibliography On

Exclude matches

< 3 words